Bounded Rationality in Pricing under State Dependent Demand: Do Firms Look Ahead? How Far Ahead?*

Hai Che
Walter A. Haas School of Business
University of California at Berkeley
545 Student Services Building
Berkeley, CA 94720-1900
Email: haiche@berkeley.edu
Phone: (510) 643-8918

K. Sudhir
Yale School of Management
135 Prospect St, PO Box 208200
New Haven, CT 06520
Email: k.sudhir@yale.edu
Phone: (203) 432-3289

P.B. Seetharaman
Jesse H. Jones Graduate School of Management
Rice University
P.O. Box 2932
Houston, TX 77252-2932
Email: seethu@rice.edu
Phone: (713) 348-6342

This Draft: July 2006

* This paper is based on Chapter 1 of the first author's doctoral dissertation at Washington University. We thank Chakravarthi Narasimhan, Tat Chan, Chuck Moul, Miguel Villas-Boas and Pradeep Chintagunta for feedback and detailed comments. The authors also thank participants at the Marketing workshops at Berkeley, Carnegie Mellon, Chicago, Houston, INSEAD, Minnesota, National University of Singapore, NYU, Rice, Rochester, Rutgers, Stanford and SUNY-Buffalo for comments.

Bounded Rationality in Pricing under State Dependent Demand:
Do Firms Look Ahead? How Far Ahead?

Abstract

The authors propose an empirical procedure to investigate the pricing behavior of manufacturers and retailers in the presence of state dependent demand. Rather than *assuming* that firms are perfectly forward looking and accordingly solving for dynamic equilibria that would arise in the presence of state dependence, the authors systematically evaluate whether boundedly rational firms indeed look ahead and if so, to what extent when setting prices. They illustrate the procedure using household-level scanner panel data on breakfast cereals, and replicate the substantive results using data on ketchup. The authors find that (1) Omission of state dependence in demand biases inference of firm behavior, i.e., one erroneously infers tacit collusion, when firms are competitive; (2) The observed retail prices are consistent with a pricing model in which both manufacturers and retailers are forward looking, i.e., they incorporate the effects of their current prices on their future profits. However, the firms' time horizon when setting prices is short-term, i.e., they look ahead by only one period, suggesting that firms are boundedly rational in their dynamic pricing behavior; (3) Even a myopic pricing model of firms that accounts for state dependence in demand is a reasonable approximation of the observed prices in the market.

Keywords: Brand Choice, State Dependence, Inertia, Variety-Seeking, Pricing, Dynamic Pricing, Bounded Rationality, Scanner Panel Data, Empirical Industrial Organization.

Bounded Rationality in Pricing under State Dependent Demand:
Do Firms Look Ahead? How Far Ahead?

Since Guadagni and Little's (1983) pioneering application illustrating the estimation of a brand choice model on household scanner panel data, a rich literature in marketing has documented that state dependence is an important driver of households' brand choices in several frequently purchased, packaged goods categories. State dependence arises due to two behavioral phenomena: inertia and variety-seeking. While inertia refers to a household repeat-purchasing a brand on account of habits formed out of past consumption experiences, variety-seeking refers to a household switching from one brand to another on account of satiation with attributes consumed in the past.

A key empirical generalization of this literature is that state dependence effects are distinct and separate from the effects of unobserved heterogeneity across households in their brand choice probabilities. Keane (1997) shows that unobserved heterogeneity is overestimated when state dependence is omitted. Further, disentangling state dependence and heterogeneity is critical to accurately infer price elasticities (e.g., Roy, Chintagunta and Haldar 1996; Seetharaman 2004) and competitive market structure (Erdem 1996). In sum, accounting for state dependence is critical to make an accurate inference of market demand.

Given the prevalence of state dependent demand in many product categories, a pertinent question that arises is whether firms – manufacturers and retailers – account for such state dependencies in demand while setting prices for their brands. For example, do firms account for the effects of lagged demand while setting current prices for their brands? Further, do firms account for the effects of current prices on future profits while setting these current prices, i,e, are they forward looking when setting prices? If so, how far in to the future do firms look ahead? What are the research consequences to the econometrician of ignoring such firm behavior while estimating supply-side pricing models? Answering these questions, that have been ignored in the existing literature on state dependence, is of critical research and policy interest. Our purpose in this study is to propose and illustrate an empirical procedure that allows us to answer such questions.

In recent years, an emerging literature, grounded in Empirical Industrial Organization, has focused on inferring supply-side strategic behavior among firms using data on weekly marketing activities of firms and aggregate (i.e., market-level or store-level) demand data on various brands (e.g., Kadiyali 1996; Roy, Hanssens and Raju 1994; Sudhir 2001b). We employ these methods, appropriately extending them to handle inter-temporal pricing behavior of firms, to address the questions raised in the previous paragraph. Since it is extremely difficult, if not impossible, to disentangle state dependence from unobserved heterogeneity using aggregate (market- or store-level) demand data (for an early illustration, see Jeuland 1979), we employ disaggregate (household-level) demand data for such purpose. First, we test how ignoring state dependence affects the inference of competitive pricing behavior of firms in the market. We then address the following questions about firms' pricing behavior.

1. To what extent do firms take inter-temporal effects of state dependence into account in setting prices for their brands?

When households' brand choices are state dependent, current brand choices will affect future brand choices. Therefore, firms should normatively take into account the impact on future demand when setting current prices. For example, if there is inertia, firms may lower current prices to lock in customers for the future (Klemperer 1987; Villas-Boas 2004; Freimer and Horsky 2004). But past research has found that managers find it difficult to correctly accommodate inter-temporal effects in their decision-making (e.g., Chakravarti, Mitchell and Staelin 1979, Meyer and Hutchinson 2001), but do better when taking into account merely current effects (McIntyre 1982). Hence simpler "shorter-horizon" games may explain firm behavior better than more complex "longer-horizon" games. It is, therefore, an empirical question as to the extent to which firms account for inter-temporal effects of state dependence (i.e., how far do they look ahead?) in setting prices (see Hoch, Camerer and Lin 2006 for insightful discussions on bounded rationality in firm behavior).

2. What is the relative gain from modeling forward looking behavior by firms, relative to a myopic pricing model, when consumers' brand choices are state dependent?

A myopic pricing model, when consumers' brand choices are state dependent, takes into account the effects of previous brand choices of consumers on current price elasticities and, therefore, current optimal prices of brands. However, unlike a forward looking model, the myopic model does not take into account the effects of current brand choices of consumers on future price elasticities and, therefore, future optimal prices of brands. A pertinent question that arises, then, is how much do firms account for one effect ("effect of the past") versus the other ("effect on the future") while setting current prices for their brands? It is possible that the myopic pricing model captures sufficient variation in observed prices of brands that it yields reasonable conditional forecasts of competitors' prices even though it ignores forward looking behavior of firms. Such a finding would be of practical value to firms since the computational burden associated with solving dynamic Nash equilibria -- that arise in inter-temporal pricing games played by competitive firms – is significantly higher than that associated with solving a myopic pricing model.

The rest of the paper is organized as follows. Next, we present the proposed models of demand and supply, as well as the estimation procedure. We then discuss the data and empirical results. Finally we conclude with a summary of the key findings, and a discussion of limitations and future research.

Model Development

We model a market with multiple manufacturers who sell to consumers through multiple retailers. We follow the previous literature in assuming that each retailer is a local monopolist (e.g., Besanko, Gupta and Jain 1998; Sudhir 2001b, Villas-Boas and Zhao 2005). This assumption is justified for many packaged goods categories because estimated cross-store elasticities are fairly small in magnitude (Walters 1991). We present our model in three parts: (1) Demand Model (2) Myopic Supply Model and (3) Forward-Looking Supply Model.

Demand Model

We model demand using a household-level multinomial logit model. The probability of a household h $(h=1,2,\ldots,H)$ purchasing one of J available brands (denoted by $j=1,\ldots,J$)[1] or not

[1]For model exposition, we will assume that all J brands are always available, but brand unavailability is taken into account in the estimation.

purchasing in the category (j=0) on a shopping occasion $t=1,2,\ldots,T_h$ is given by:

$$\theta_{hjt} = \frac{\exp(v_{hjt})}{\sum_{k=0}^{J}\exp(v_{hkt})} \tag{1}$$

where v_{hjt} is given by $v_{hjt} = X_{hjt}\cdot\beta_h - \alpha_h\cdot p_{hjt} + SD_h\cdot sim_{ij} + \xi_{jt}$ for brand j ($j=1,\ldots,J$), and is given by $v_{h0t} = \delta_h\cdot INV_{ht}$ for the outside good. The vector X_{hjt} denotes the set of explanatory variables characterizing brand j, as experienced by household h, on shopping occasion t, while β_h denotes the corresponding set of response coefficients. The vector X_{hjt} includes brand characteristics and marketing mix variables – specifically feature and display – excluding price. p_{hjt} denotes the retail price of brand j at shopping occasion t, while α_h denotes the corresponding price coefficient. INV_{ht} denotes the product inventory held by household h at shopping occasion t (which, in the usual manner, is imputed based on the household's observed total purchase quantity over the study period and by assuming a constant weekly consumption rate, and mean-centered within household), while δ_h denotes the corresponding inventory coefficient. ξ_{jt} denotes a composite (stochastic) measure of unobserved (to the econometrician) characteristics of brand j at shopping occasion t, which is common across all households. It refers to common demand shocks that affect all households (such as brands' national advertising campaigns on television, macro-economic conditions etc. that are not recorded in scanner panel data and, therefore, unobservable to researchers, but are observable by the price-setting firms).

sim_{ij} is a similarity variable that captures how much perceived similarity brand j has to the brand i bought by household h on its previous purchase occasion, and is operationalized as $sim_{ij}=(I_{ij}+\sum_l r_l\cdot I_{ijl})/(1+\sum_l r_l)$, $l=1,\ldots,L$, where i refers to the brand purchased by household h at its previous purchase occasion, L stands for the number of product attributes represented among all brands within the product category. I_{ij} is an indicator variable that takes the value 1 if $i=j$, i.e., i and j are the same brand, and 0 otherwise, I_{ijl} is an indicator variable that takes the value 1 if brands i and j share the same level of attribute l and 0 otherwise (for example, if brand i is high-sugar, and brand j is low-sugar, they will be assumed not to share the same level of the attribute sugar; on the other hand, if both brands are high/low on fiber, they will be assumed to share the same level of the attribute fiber), and $r_l>0$ stands for the perceived importance for attribute l in determining inter-brand similarity. Note that sim_{ij} is restricted to lie between 0 and 1, and is monotonically increasing in the number of attributes shared by brands i and j. This specification enables us to separately identify sim_{ij} from the *unobserved* state dependence parameter SD_h, which is unrestricted in magnitude. Since sim_{ij} is defined relative to the brand i chosen by the household at its previous purchase occasion, its value remains unchanged from one shopping trip to the next if the household does not make a purchase in the product category.

We will use a couple of examples to illustrate how the similarity variable sim_{ij} works. Suppose household h had purchased brand k at its previous purchase occasion. For this household, the similarity variable for brand k at the current shopping trip is $sim_{kk}=(I_{kk}+\sum_l r_l \cdot I_{kkl})/(1+\sum_l r_l)= (1+\sum_l r_l \cdot 1)/(1+\sum_l r_l)=1$. For the same household, the similarity variable for brand m -- which shares no attributes whatsoever with brand k -- at the current shopping trip is $sim_{km}=(I_{km}+\sum_l r_l \cdot I_{kml})/(1+\sum_l r_l)= (0+\sum_l r_l \cdot 0)/(1+\sum_l r_l)=0$.

The effect of how similarly brand j is perceived in comparison to the brand i purchased by household h at its previous purchase occasion, i.e., sim_{ij}, can be positive or negative, depending on whether state dependence effects are positive or negative respectively, which are captured by the state dependence coefficient SD_h. For example, if a household is inertial, $SD_h>0$, and the effect of purchasing a brand similar to the one previously purchased will be to increase utility. Conversely, if a household is variety-seeking, $SD_h<0$, and the effect of purchasing a brand similar to the one previously purchased will be to decrease utility. We allow the state dependence parameter, SD_h, to be a function of household-specific demographic variables: $SD_h=\gamma_{oh}+DEMO_h \cdot \gamma_{1h}$, where $DEMO_h$ is a vector of demographic variables (such as family size, income etc.) characterizing household h, γ_{1h} is the corresponding vector of parameters, and γ_{oh} is the baseline state dependence level of households in the market. This allows the estimated degree of inertia or variety-seeking to be an explicit function of household characteristics. We accommodate unobserved heterogeneity across households by allowing household-specific parameters to follow a latent class distribution across households (Kamakura and Russell 1989).

The market share of brand j at period t is given by $S_{jt}=\sum_h \theta_{hjt} /H,\ h=1,\ldots,H$ where θ_{hjt} is the *unconditional* (i.e., integrated over the unobserved heterogeneity distribution) brand choice probability of household h, taking its lagged brand choice i as given.

Myopic Supply Model

We develop a supply model where firms set prices only taking into account current demand (and its dependence on past demand through state dependence) and cost conditions. Since we wish to empirically infer the nature of *horizontal strategic interactions* between manufacturers and *vertical strategic interaction* between each manufacturer and the retailer, we estimate a menu of games allowing for different combinations of horizontal and vertical strategic interactions and pick the best fitting model. We allow for two types of horizontal interactions between manufacturers: (1) Bertrand, and (2) Collusive (Note: We also test for levels of cooperative behavior between manufacturers that are intermediate between Bertrand and Collusive, using the weighted objective function approach of Sudhir 2001a); and two types of vertical strategic interactions between the manufacturers and the retailer: (1) Manufacturer Stackelberg, and (2) Vertical Nash.

The retailer sets the vector of retail prices $p_t=(p_{1t},\ldots,p_{Jt})$ in week t to maximize category profits, i.e., $\max \pi_R(p_t)=\sum_j(p_{jt}-w_{jt})S_{jt}(p_t)M,\ j=1,\ldots,J$ where J is the number of brands in the category, w_{jt} is the wholesale price of brand j at time t, $S_{jt}(p_t)$ is the market share of brand j at time t, and M is the size of the retailer's local market. We follow the previous literature in this area in assuming that the retailer is a monopolist (e.g., Besanko, Gupta and Jain 1998; Sudhir 2001b), because cross-elasticities across retailers have been found to be very low (e.g., Walters 1991). Berto Villas-Boas (2004) models cross-store competition, but finds very little evidence of cross-store competition. We assume that manufacturers offer different wholesale prices to different retailers. Given that retailers negotiate discounts individually with manufacturers, this is a reasonable assumption.

Suppose there are F firms, each of which produces some subset, F_f, of the $j=1,\ldots,J$ different brands in the product category. A manufacturer pricing under Bertrand competition sets wholesale prices to maximize total profits from its product line as follows: $\pi_f=\sum_{F_f}(w_{jt}-mc_{jt})S_{jt}(p_t)M$ where mc_{jt} is the marginal cost of producing brand j, and the other terms are as explained earlier. In contrast, a manufacturer pricing under collusion sets wholesale prices to maximize total profits from all products sold as follows. $\pi_f=\sum_j(w_{jt}-mc_{jt})S_{jt}(p_t)M,\ j=1,\ldots,J$. Following the procedures developed in Sudhir (2001b) and Berto Villas-Boas (2004), we solve the first order conditions of the retailer and manufacturer to obtain the econometric pricing equation below. The derivation is provided in the appendix.

$$
\begin{aligned}
p_t &= mc_t + (w_t - mc_t) + (p_t - w_t) \\
&= \underbrace{mc_t}_{\text{Marginal Cost}} \underbrace{-(\Phi_t' G_t^{-1} \Omega_t)^{-1} S(p_t)}_{\text{Manufacturer Margin}} + \underbrace{\Phi_t^{-1} S(p_t)}_{\text{Retail Margin}}
\end{aligned}
\qquad (2)
$$

For the Manufacturer Stackelberg Model:

(1) Φ_t is a $J \times J$ matrix with elements $\Phi_{jrt} = -\frac{\partial S_{rt}(p_t)}{\partial p_{jt}}, j, r = 1, \ldots, J$;

(2) $$G_t = \begin{bmatrix} \left(2\frac{\partial S_{1t}}{\partial p_{1t}} + \sum_{k=1}^{J}(p_{kt} - w_{kt})\frac{\partial^2 S_{kt}}{\partial p_1^{2t}}\right) & \cdots & \left(\frac{\partial S_{Jt}}{\partial p_{1t}} + \frac{\partial S_{1t}}{\partial p_{Jt}} + \sum_{k=1}^{J}(p_{kt} - w_{kt})\frac{\partial^2 S_{kt}}{\partial p_{Jt}\partial p_{1t}}\right) \\ \vdots & \ddots & \vdots \\ \left(\frac{\partial S_{1t}}{\partial p_{Jt}} + \frac{\partial S_{Jt}}{\partial p_{1t}} + \sum_{k=1}^{J}(p_{kt} - w_{kt})\frac{\partial^2 S_{kt}}{\partial p_{1t}\partial p_{Jt}}\right) & \cdots & \left(2\frac{\partial S_{Jt}}{\partial p_{Jt}} + \sum_{k=1}^{J}(p_{kt} - w_{kt})\frac{\partial^2 S_{kt}}{\partial p_{Jt}^2}\right) \end{bmatrix}$$

and (3) Ω_t is a J by J matrix with $\Omega_{jrt} = \Omega^*_{jr} \times \Phi_{jrt}$, where $\Omega^*_{jr} = \begin{cases} 1, & if\ \exists f : \{r, j\} \subset F_f \\ 0, & otherwise \end{cases}$ for the Bertrand Equilibrium and Ω^* is a J by J square matrix with all the elements being 1 for the collusive equilibrium.

For the Vertical Nash Model, we replace $\Phi_t' G_t^{-1}$ by $-I$, a J-dimensional identity matrix, and retain the rest of the equations as above.

Forward-Looking Supply Model

We estimate the same set of models of horizontal and vertical strategic interactions as we did under the myopic model. But in the forward-looking model, retailers and manufacturers look ahead and base their pricing decisions on their impact on not only current sales, but also future sales. To assess how far firms look ahead, we sequentially estimate models where firms look ahead T periods, $T=1,2,3,\ldots$ and stop when there is no improvement in model fit going from $T=\tau$ to $T=\tau+1$. Here we present the estimation equation for a 1-period look-ahead model, and provide the derivation in the appendix. The estimation equations for the general T-period look-ahead model, for $T>1$, are in a technical appendix that is available from the authors.

The retailer's objective in the 1-period look-ahead model is to set prices such that they maximize category profits from the current period and the next period (discounted by a factor δ), i.e., $\max V_R=\pi_R(p_t)+\delta\pi_R(p_{t+1})$. Similarly, the manufacturer *f*'s objective is to maximize $\max V_f=\pi_f(w_t)+\delta\pi_f(w_{t+1})$. Since the second-period (i.e., next period) prices are set without look ahead, the second-period profit margin will be the same as that derived under the myopic pricing model. But the first-period (i.e., current period) margins will change to reflect the look-ahead. The supply side pricing equation can then be decomposed as follows:

$$\begin{aligned} p_t = \underbrace{mc_t}_{\text{Marginal Cost}} &+ \underbrace{PCM_M^{myopic}(t=1) - \delta\Delta_1\left[PCM_M^{myopic}(t=2)\right]}_{\text{Manufacturer Margin}} \\ &+ \underbrace{PCM_R^{myopic}(t=1) - \delta\Delta_1\left[PCM_R^{myopic}(t=2)\right]}_{\text{Retail Margin}} \end{aligned} \tag{3}$$

where $\Delta_1 = \bar{\theta}_{r2r1} - \sum_{k=1,k\neq r}^{J}\bar{\theta}_{r2k1}$, $\bar{\theta}_{r2r1}$ is the unconditional (i.e., integrated over the unobserved heterogeneity distribution) transition probability for a household that *bought* brand *r* in period 1 *continuing to buy* the same brand *r* during period 2, and $\bar{\theta}_{r2k1}$ is the unconditional transition probability for a household that *did not buy* brand *r* in period 1 *switching to buy* brand *r* during period 2.

If households are inertial (*SD*>0) and $\bar{\theta}_{r2r1} - \sum_{k=1,k\neq r}^{J}\bar{\theta}_{r2k1} > 0$, it can be shown that retailers and manufacturers will price lower in the first-period than under the myopic case. The intuition for this analytical result (similar to Klemperer 1987 and Villas-Boas 2004) is that firms have an incentive to lower their current prices in order to *lock-in* customers for the second period. This incentive to lower price will be moderated by the number of customers already *locked in* to the firm's brand at the beginning of the first period, i.e., the inertials who bought the firm's brand during their previous purchase. However, there is another effect. If households are inertial (*SD*>0) but $\bar{\theta}_{r2r1} - \sum_{k=1,k\neq r}^{J}\bar{\theta}_{r2k1} < 0$, retailers and manufacturers will price higher in the first-period than under the myopic case. This will happen when the effects of inertia are small relative to the effects of marketing activities. This does not invalidate Klemperer's (1987) result, since he assumes a duopoly pricing game, using a Hotelling line, where the effects of switching costs on consumer demand are *always* high enough for the two firms to compete more aggressively in the first period. While in the case of oligopoly pricing and multinomial logit demand, as in our case, the effects of marketing activities on households' brand choices could dominate the effects of inertia, i.e., $\overline{\theta}_{r2r1} < \sum_{k=1,k\neq r}^{J} \overline{\theta}_{r2k1}$.

On the other hand, if households are variety-seeking (*SD*<0) and $\overline{\theta}_{r2r1} - \sum_{k=1,k\neq r}^{J} \overline{\theta}_{r2k1} < 0$, forward looking manufacturers and retailers will price higher than under the myopic case. Not only do firms not have lock-in incentives to lower prices, but also they can raise prices because variety-seeking customers who switch due to higher current prices are more likely to switch back. The competitive effects, through marketing mix activities, only strengthen the effects of variety seeking on pricing in this case, and the first-period prices will be strictly higher than under the myopic case. Since Klemperer (1987) and Villas-Boas (2004) did not model variety-seeking in their work, this is a new analytical insight in the literature on state dependence.

When there are heterogeneous households in the market, the actual prices offered by firms in each period will be based on the composition of state-dependent households. In other words, the prices will respond to not only whether inertial or variety-seeking households are expected to arrive at the store on a given week, but also the degree of heterogeneity across these households in terms of both the degree of state dependence and the identity of the most recently purchased brand. This can be one explanation for time-varying prices observed in this product category, and is behaviorally consistent with Freimer and Horsky's (2004) normative predictions about alternating price promotions by manufacturers in consecutive periods when there is a high degree of brand choice inertia. Another source of inter-temporal dynamics in our model is the household's product inventory which influences the household's utility for the no-purchase option. To the extent that a household's current purchase decision can, therefore, affect the household's future inventory of the product, the forward looking model additionally accounts for the effects of households' inventory dynamics on firms' pricing decisions.

Estimation

We employ a two-stage estimation approach (Newey and McFadden 1994). In the first step, we estimate the demand model using the Maximum Likelihood technique, correcting for potential price endogeneity. While these estimates are unbiased and consistent, they are inefficient since the information in supply equations and, therefore, price data, are not exploited. However, the advantage is that demand estimates are unbiased by any potential mis-specification in supply equations (especially since our intent is to identify the best supply-side specification). Further, due to the large number of observations in household-level data, efficiency concerns are not as high as with aggregate data. In the second step, conditional on the demand estimates, we compute the margins under alternative supply-side specifications. We then estimate a cost function, sequentially for each of these computed margins, and choose the best fitting supply-side model on the basis of fit.

Demand-Side Estimation

Recall that a household h's probability of choosing brand j at shopping occasion t, θ_{hjt}, is a function of ξ_{jt}, which is the common shock among households that is unobserved by the econometrician. Villas-Boas and Winer (1999) argue that since profit-maximizing firms take ξ_{jt} into account when setting prices, there is a price endogeneity problem. We, therefore, need instruments for price in order to obtain an unbiased estimate of the price coefficient. Denoting $P_{jt}^{\text{Instrument}}$ as instruments for price, we can estimate the pricing equation: $P_{jt} = \gamma_0 + \gamma_1 \cdot P_{jt}^{\text{Instrument}} + \eta_{jt}$. The random errors, η_{jt}, could arise from both cost shocks and demand shocks. Therefore, when prices are endogenous, ξ_{jt} and η_{jt} should be correlated. Assuming a joint normal distribution of ξ_{jt} and η_{jt}, and $E(\xi_{jt}\eta_{jt})=\rho_j\sigma_{\xi jt}\sigma_{\eta jt}$, testing for price endogeneity is equivalent to estimating whether $\rho_j=0$ or not.

We estimate the demand model using Limited Information Maximum Likelihood (LIML). The joint likelihood of the household h over the demand and pricing equation is:

$$L_h(\alpha_h, \beta_h, \delta_h, \gamma_0, \gamma_1, \sigma_\xi, \sigma_\eta) = \prod_{t=1}^{T_h} \int \prod_{j=0}^{J} \left[\theta_{hjt}\left(\xi_{jt}\right)\right]^{y_{hjt}} \cdot f\left(\xi_{jt} \mid \eta_{jt}\right) \cdot f(\eta_{jt}) d\xi_{jt} \qquad (4)$$

where $y_{hjt} \begin{cases} =1 \text{ if } j \text{ is chosen by } h \text{ at } t \\ =0 \text{ otherwise} \end{cases}$ and $f(\eta_{jt})$ refers to the density of the distribution of η_{jt} evaluated at the estimated residual from the pricing equation estimation. We note that since scanner panel data are used in the estimation, one explicitly observes the number of no purchases (i.e., purchases of the outside good) during each week instead of having to impute them using some assumed market size.

During estimation, we evaluate the integral contained in the likelihood function above using simulation. Previous studies applying LIML (such as Villas-Boas and Winer 1999) do not allow for heterogeneous demand parameters due to the computational burden involved. In contrast, we allow for different segments of households that differ from each other not only in their propensity for variety-seeking or inertia, but also in their responsiveness to the marketing mix and intrinsic preferences for the large number of choice alternatives (>100). We estimate a latent class model of unobserved heterogeneity (Kamakura and Russell 1989), under which the sample likelihood function becomes

$$L = \prod_{h=1}^{H} \left\{ \prod_{k=1}^{K} \left[L_{hk} * \Pr(k) \right] \right\} \tag{5}$$

where K stands for the number of supports of the discrete heterogeneity distribution, L_{hk} stands for the likelihood function computed for household h under the assumption that h belongs to segment k (and is evaluated as in equation 4), and $\Pr(k)$ stands for the (prior) probability of household h belonging to segment k.

Supply-Side Estimation

We assume that the marginal cost for each brand j, mc_{jt}, is a brand-specific linear function of observable cost shifters, as shown below.

$$mc_{jt} = mc_{j0} + \sum_{k} input_{jkt} \cdot \omega_k \tag{6}$$

where mc_{j0} is a brand-specific baseline marginal cost, $\omega = (\omega_1, \omega_2, \ldots, \omega_K)$ is a vector of coefficients of *brand-specific* input prices, including the *manufacturer-specific* wage rate, at time t ($input_{jkt}$). We compute the manufacturer and retailer margins ($\widehat{PCM}_{\mathrm{M},sjt}$ and $\widehat{PCM}_{\mathrm{R},sjt}$ respectively) conditional on alternative supply specifications, based on the demand estimates. Subtracting out these computed margins from prices, we then formulate the cost equation as follows:

$$p_{sjt} - \widehat{PCM}_{M,sjt} - \widehat{PCM}_{R,sjt} = \gamma_0 + \gamma_s \cdot \text{Store}_s + \gamma_m \cdot \text{Manufacturer}_j + \sum_k input_{jkt} \cdot \omega_k + \eta_{sjt} \qquad (7)$$

where η_{sjt} is a random error, which is assumed to be normally distributed with mean zero and covariance matrix V_η.

In order to estimate the cost equation, we write the empirical likelihood function as the likelihood of observed prices given the distribution assumptions on η_{sjt}. Denoting $g(.)$ as the marginal (normal) density of η_{sjt} , this likelihood function is written as:

$$L = \prod_{t=1}^{T} \left[\prod_{j=1}^{J} g(\eta_{sjt}) \right] \qquad (8)$$

The 2-period supply-side model is estimated as follows (the same backward-induction procedure is used to estimate 3- and 4-period supply models): 1) Using the demand-side estimates, calculate the margin for each brand for weeks 2 to 104 (using equation 2); 2) Treating the computed margins from the previous step as second-period margins of two-period pricing problems, compute the corresponding first-period margins for weeks 1 to 103 (using equation 3); 3) Estimate the first-period pricing equations using price data for weeks 1-103 (using equations 7 and 8). We use a simulation (details available in a technical appendix), to verify that our proposed estimation approach indeed produces consistent estimates of model parameters.

Data

We need household-level data, as opposed to aggregate data, in order to identify state dependence. We use IRI scanner panel data in the breakfast cereal category, drawn from urban and suburban areas in a large U.S. city, covering a 104-week period from June 1991 to June 1993. In the suburban market, 483 households purchase 24,971 units (measured in category-specific regular package sizes) of cereals in 68,432 shopping trips, while in the urban market, 480 households purchase 23,011 units of cereals in 71,247 shopping trips[2]. The purchases are made from stores belonging to different chains.

[2]We estimate the demand models separately for these two markets. We report only the estimation results for the suburban market. Results for the urban market are available from the authors.

In all, these households purchase 117 brands. If households buy cereal on different days of the same week, we treat each purchase separately for estimation. The data include information about households' demographics, such as family size, income, household head's age, kids' ages, etc.

Since we model state dependence in attribute space, we require data on product attributes. We collected such information by visiting supermarkets and reading product labels, and also from manufacturers' websites. When brands had been discontinued, we collected attribute information from cereal reviews found in various sources.

We collected weekly cost data for different ingredients (corn, wheat, rice and oatmeal), sugar, fruits (apple, grape) from the Department of Agriculture and weekly labor wage data for 5 different states (corresponding to locations of cereal manufacturer plants) from the U.S. Current Population Survey Annual Earnings File. A brand's cost is based on its ingredients and state of production. For example, Rice Krispies' input cost is based on the cost of rice and sugar and Michigan's labor wage rate. We collect information on other production and packaging costs (gasoline and "folding paperboard boxes, including retail food") from the Producer Price Index.

Estimation Results

We report the results of the empirical analysis in two parts: (1) Tables 1 - 2 report the demand-side results,[3] and (2) Tables 3 - 5 report the supply-side results.

Demand-Side Results

We estimate heterogeneity in a latent class framework, by increasing the number of supports for the heterogeneity distribution until there is no improvement in fit. To assess the role of state dependence, we evaluate three models: (1) a zero-order multinomial logit model with no state dependence (McFadden 1980), (2) a multinomial logit model with a lagged brand choice dummy to capture state dependence (Seetharaman, Ainslie and Chintagunta 1999), and (3) a multinomial logit

[3]The estimation was done in the C++ programming environment on a personal computer employing a Pentium 2 GHz chip. Convergence times varied from one day (for the static models without endogeneity correction) to one week (for inter-temporal models with endogeneity correction).

model with the proposed attribute-based similarity specification of state dependence. We find that the three-support heterogeneous version of the proposed attribute-based state dependence model has the best fit (see Table 1), in terms of both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). This demonstrates that the proposed state dependence model is superior to extant approaches of modeling state dependence. It also suggests the importance of modeling demand at an SKU level in categories where a brand has a number of SKUs with widely varying attributes (as argued by Fader and Hardie 1996 and Ho and Chong 2003).

As in previous research (Keane 1997; Seetharaman, Ainslie and Chintagunta 1999), we find that the magnitude of the estimated price coefficient at the market level (which is a weighted average of support-specific price coefficients), α, as well as all the price elasticities, are overstated when state dependence effects are ignored, i.e., $|\alpha^{\text{No SD}}| > |\alpha^{\text{With SD}}|$ (see Table 1). This suggests that the optimal Bertrand-Nash prices will be systematically *lower* if one bases it on estimates of demand models that ignore state dependence. We revisit this issue when we discuss the supply-side estimates.

We further discuss only the results of the three-support heterogeneous version of the attribute-based state dependence model, since it has the best fit. Table 2 contains the estimates of this model with and without price endogeneity accounted for. As in Nevo (2001), we use prices from another market as instruments to account for price endogeneity. However, these instruments will only partially correct the endogeneity bias if national advertising creates common demand shocks across markets. Unfortunately, instruments for prices used by other researchers do not work in our context. Lagged prices are not valid instruments because common unobservables affect both current prices and future prices through state dependence. Attribute-based instruments also do not work because product attributes remain unchanged during the period of analysis, providing no variation in these instruments. We acknowledge that displays and features may be potentially endogenous even though Sudhir (2001b) rejects feature endogeneity once price endogeneity has been accounted for. Since our focus is on prices and we do not have good instruments for features and displays, we abstract away from the potential endogeneity of other variables in the paper. The R^2 of the pricing regression with other market price (instrument) is .52, and the F-statistic is highly significant (F (1, 36,503) =39113).

The Hausman (Hausman and McFadden 1984) test rejects the null hypothesis that prices are exogenous. As expected, the magnitude of the estimated price coefficient increases with endogeneity correction. Interestingly, including state dependence *reduces* the magnitude of the estimated price coefficient, while correcting for endogeneity *increases* it. The remaining coefficients remain fairly similar between the two models.

We uncover three segments of sizes 27%, 16% and 57%. The largest support is extremely price sensitive, while the remaining two supports are relatively price insensitive. We report aggregate summary statistics for *SD* (e.g., mean, median) within each of the three supports. To the extent that supports 1, 2 and 3 are observed to have negative (average estimated *SD* of -1.35), positive (average estimated *SD* of 5.08) and insignificant state dependence coefficients respectively, we will refer to them as the variety seeking, inertial and zero-order segments respectively. Thus we empirically document the existence of a variety seeking segment in the market using scanner panel data for the **first** time. This is in contrast to previous research (e.g., Erdem 1996, Seetharaman, Ainslie and Chintagunta 1999), who find that, almost without exception, households tend to be inertial.

We profile each of the uncovered segments in terms of the average demographic characteristics of members belonging to the segment. From the profiling results, it is clear that variety-seeking households are, on average, larger than inertial households. This is consistent with the intuition that variety-seeking is partly an artifact of aggregating brand choices made by household members who are heterogeneous in their tastes. Technically, this should be considered as intra-household heterogeneity in tastes and not variety-seeking (Kahn, Kalwani and Morrison 1986; Lee, Sudhir and Steckel 2002). However, since we do not observe consumption, it is impossible to disentangle intra-household heterogeneity and variety seeking (as with all previous research in this area). We explore this issue in more detail by restricting attention to single member households.

We find some differences in the estimated relationships between demographics and state dependence across the different segments. The sign on the family size variable is positive within segment 1, but negative within segment 2, which suggests an interesting second-order effect, i.e., within each segment, larger families show less state dependence (i.e., more zero-order behavior) than smaller families. To the extent that previous studies have restricted demographic effects to be equal across segments (not expecting them, a priori, to be different), such second-order effects have not been estimated before. However, we find that zero-order households, i.e., those in segment 3, become more inertial as their family size increases. The profiling also shows that, on average, variety-seeking households have higher-income than inertial households, but as income increases within each segment, households are found to become less state dependent in their brand choices. Again, zero-order households of segment 3 behave differently compared to this, by becoming more inertial as their income increases. Further research (possibly using laboratory experiments) is required to understand the behavioral rationale of these curious second-order effects, which indicate non-linear dependencies of state dependence on demographics. This also raises the intriguing question of whether insignificant demographic effects estimated in previous brand choice models (that restrict demographic effects to be equal across segments) could be an artifact of such opposing effects that 'cancel each other out.'

To explore whether variety-seeking effects survive beyond intra-household heterogeneity, we re-estimate our proposed brand choice model using purchase data of *single-member households* only (205 out of 963 households). We still uncover a segment of variety-seekers, although the size of the variety-seeking segment is smaller (10.2%). To the extent that the estimated degree of variety-seeking in this market is only mitigated, but not eliminated, in an analysis of single-member households, we believe that variety-seeking indeed characterizes brand choices of some households in the breakfast cereal category.

A key element of our proposed brand choice model is that state dependence in households' brand choices over time is driven by observed product attributes. In other words, inertial households choose brands with similar attributes, while variety-seeking households choose brands with dissimilar attributes, on successive purchase occasions. We see that sugar-level, fruit-nut and product-type (kids, family, adult/family) all have positive and significant effects on perceived inter-product similarity between alternatives. Ingredients (corn, wheat, rice and oatmeal) and fiber do not significantly affect perceived inter-product similarity. These findings suggest that promotional messages by marketers using easily communicable taste characteristics, such as sugar level and the presence of fruits and nuts, will be more effective in influencing perceptions and choices of households than using less easy-to-discern characteristics, such as ingredients or fiber. For example, Kellogg's could target variety-seeking households (if they can be identified using appropriate data sources) with a portfolio of coupons across brands that are diverse on discernible product attributes to ensure that variety-seeking households stay within Kellogg's franchise.

In terms of household preferences for attributes, we see that variety-seekers prefer high-fiber and fruit-nut cereals to sweetened cereals, while inertials prefer high-fiber and zero-order households prefer sweetened and fruit-nut cereals. Overall, product attributes are observed to significantly explain households' brand preferences in this product category. Both inertial and variety-seeking segments of households exhibit lower price sensitivity than the zero-order segment, and also have positive coefficients for feature and display. Product inventory increases the utility of the outside good and reduces purchase incidence only for variety seekers; for the other two segments, it has no significant effect. The behavioral literature (e.g., McAlister 1985) has shown that variety seeking households satiate on the product attributes in existing product inventory, and switch brands accordingly; our findings suggest that variety seeking households also respond strongly to product inventory in their category purchase decisions.

We calculate price elasticities of different brands based on the estimated demand parameters. Cheerios, a large share brand in the product category, has an own-price elasticity of -1.91. However, Quaker Oats, a similar brand in terms of product attributes, but with much smaller market share, has a much higher own-price elasticity of -7.8. In fact, large share brands are observed to have smaller own-price as well as cross-price elasticities compared to small share brands, which is consistent with their strong market positions. This shows that the estimated extent of unobserved heterogeneity across households is strong enough to relax the restrictive IIA property of the homogeneous logit specification, which would imply larger price elasticities for larger share brands. We uncover a systematic reduction in price elasticities after we account for state dependence. On average, the two elasticities differ by 10-20%. For example, Cheerios' own price elasticity is estimated to be -2.15 if one ignored state dependence, compared to -1.91 after accounting for state dependence; the difference is statistically significant at the 5% level.

Supply-Side Results

We estimate three types of pricing games: static (i.e., ignoring state dependence in demand), myopic (i.e., accommodating state dependence in demand, but ignoring inter-temporal behavior of firms), and inter-temporal (i.e., accommodating both state dependence in demand and 1-period look ahead). For each type of pricing game, we estimate [2 VSIs (Vertical Nash, Stackelberg) * 2 HSIs (Bertrand, Collusive)] = 4 types of interactions between firms. Therefore, we estimate 4 (interaction types per pricing game) * 3 (pricing games) = 12 supply-side specifications, each of which would generate different predicted values of manufacturer and retailer margins. Then, based on the best-fitting model among these 12, we also estimate 2-period and 3-period look-ahead versions of the games to investigate how many periods firms look ahead into the future when pricing.

Table 3 shows the maximized log-likelihood values associated with the various supply-side specifications. We report the results for the three types of pricing games under three columns: (1) static (i.e., ignoring state dependence), (2) myopic, and (3) inter-temporal (1-period look-ahead). The Vuong test for non-nested models is used to select the best-fitting supply-side specification. The Vuong test statistics are reported below the log-likelihood values.

These results provide insights about how ignoring state dependence affects our inference of supply side pricing behavior. Consider the differences between the results for the static model (that ignores state dependence) and the myopic model (that accounts for state dependence, but does not account for inter-temporal effects). For the static model, the collusive pricing interactions between manufacturers outperform Bertrand pricing interactions, whereas the result reverses for the myopic model (that accounts for state dependence). Thus, *one could wrongly attribute the seemingly high observed prices in the cereal category to collusion between manufacturers if one ignored the effects of state dependence.* While it is well-known that demand-side mis-specification can affect inference of supply-side behavior, our results specifically show that omission of state dependence (which is the prevailing norm in the existing literature) seriously biases the supply-side inference. The reason for this finding is that ignoring the effects of state dependence in the model leads us to spuriously infer that the market is more price elastic than it truly is (i.e., the estimated price elasticities are larger in magnitude compared to their counterparts yielded by the model that includes state dependence), which puts the onus of explaining the seemingly high prices in the market on tacitly collusive pricing behavior between firms. Once we correctly accommodate the effects of state dependence in the model, the estimated price elasticities become sufficiently smaller in magnitude that the observed prices can be rationalized by Bertrand competition between firms.

We now further evaluate the supply-side results to address our two research questions about firm pricing behavior.

To what extent do firms take inter-temporal effects of state dependence into account in setting prices?

Our first research question is related to the issue of whether firms are forward looking and if so, to what extent, when setting prices. As discussed earlier, while normatively firms *should* be forward looking (e.g., Klemperer 1987), behavioral research suggests that managers have difficulty taking into account inter-temporal effects (e.g., Chakravarthi, Mitchell and Staelin 1979). We use the Vuong non-nested test to compare the goodness-of-fit between the myopic pricing model and the 1-period look-ahead inter-temporal pricing model. The test statistic is 2.75 and significant at the .01 level. We find that the 1-period look-ahead model fits the data better than the myopic pricing model, suggesting that firms -- manufacturers and retailers -- account for the effects of their current pricing decisions on next-period profits. There is little difference in terms of fit between the 1-period and 2-period look ahead models and even a reduction in model fit from the 2-period to the 3-period look-ahead model. *We conclude, therefore, that the 1-period look-ahead pricing model is an adequate characterization of inter-temporal pricing behavior of firms in this product category*.[4]

We also look at the mean square errors (MSE) between fitted and observed prices. These results are reported in Table 4. We find that the MSE is lowest for the 1-period look-ahead model that accounts for state dependence. The fit for the 2-period model is worse than the fit for the 1-period look-ahead model. Thus the one period look-ahead model provides better forecasts than the 1-period look-ahead model.

These results are consistent with the hypothesis of boundedly rational firms. In many scenarios

[4] Note that the fit improvement in going from a myopic to a forward-looking pricing model can be due to both inventory effects and state dependence. Decomposing the effects, we find that 61% of improvement is due to state dependence, while 39% is due to inventory effects. These results are available from the authors. We thank an anonymous reviewer for (implicitly) alerting us to do this decomposition.

involving dynamics, researchers typically assume that agents solve a dynamic program. But recent studies in experimental economics (Camerer, Ho and Chong 2004) suggest that even strategic players endowed with the highest thinking ability rarely think more than two steps ahead. Further, in an oligopolistic market, where the effects of price promotions on market share may cancel out very quickly, the effects of state dependence (which is a function of market share) may not extend beyond one or two periods. Our results show that it is possible that modeling shorter-term dynamics may lead to better fit and forecasts than solving a full dynamic program. Given that shorter horizon dynamic programs are computationally easier to solve, this makes it easier for firms to conditionally forecast their competitors' dynamic price responses to their chosen prices. We discuss this issue again later when discussing our simulation results.

What is the relative gain from modeling forward looking behavior by firms, relative to a myopic pricing model, when consumers' brand choices are state dependent?

Having found that firms take into account inter-temporal effects of current choices on future demand while setting current prices, we now address the question of whether a myopic model of firm behavior, when consumers' brand choices are state dependent, is a reasonable approximation of observed prices in the market. To assess this, we use two in-sample fit measures: (i) Maximized Log-Likelihood associated with the price outcomes, and (ii) Mean Squared Errors of the differences between the fitted prices and the observed prices in the data.

As far as the first in-sample fit measure is concerned, the maximized log-likelihood for (1) the best fitting static model without state dependence is 31,308, (2) the best fitting myopic model with state dependence is 33,923 and (3) the best fitting inter-temporal model is 34,456. The total percentage improvement in likelihood obtained by moving from a static model without state dependence to an inter-temporal model with state dependence is 10%. Of this 10% improvement, 8.3% comes from a myopic model with state dependence, while 1.7% is obtained by modeling the inter-temporal effects. Thus inter-temporal effects contribute to only about 17% of the improvement in likelihood, while 83% of the improvement can be obtained with simply a myopic supply model that accounts for state dependence.

As far as the second in-sample fit measure is concerned, relative to the static model without state dependence, the myopic pricing model with state dependence reduces the MSE by 22.34% (see Table 4). Further, relative to the static model without state dependence, the inter-temporal pricing model – that allows for both state dependence and forward-looking behavior -- reduces the MSE by 23.67%. The incremental reduction in MSE from accommodating forward-looking behavior, in addition to state dependence, is just 1.71%. Therefore, 94% (22.34/23.67) of the reduction in MSE is obtained by accounting for state dependence within a myopic model, while only 6% of the reduction is obtained by incorporating forward-looking behavior. Thus, our results show that the relative gain from modeling inter-temporal forward looking behavior is low relative to the myopic supply model. What is critically important to note is that the supply models indeed take into account the effects of *household purchase dynamics* through state dependence.

From a practical point of view, in conditionally forecasting competitors' prices in response to their chosen price, this implies that there is much greater “bang-for-the-buck" for a firm in computing “near-optimal" prices of their competitors by modeling demand state dependence within a myopic model than by modeling inter-temporal dynamics. This is a useful empirical insight given that the computational complexity of solving dynamic oligopolistic pricing games (with 117 brands involving 6 manufacturers and strategic retailer as in our application) is far greater compared to solving a static game.

Costs

Table 5 presents the full estimation results for the best-fitting pricing models: 1-period look ahead pricing models accounting for state dependence. The Bertrand Nash models fit the data best. General Mills has a relative cost advantage compared to other manufacturers. The retailer that stocks the largest number of cereal brands in the store (92 brands per week, on average) has a relative cost disadvantage compared to the other two retailers (83 and 87 brands per week, on average, respectively). Prices of corn, rice and oats are significant predictors in the cost equation. We do not find wages to be significant in terms of explaining weekly price variations. A variance decomposition of observed prices shows that raw materials and packaging costs contribute to 13.4%, while price-cost margins contribute to 66.5% of the variance in retail prices. The remaining 20.1% of the variance is due to unobserved cost shocks.

Interestingly, the cost estimates are identical between the 1-period and 2-period pricing models. This finding runs counter to one of the implications of the illustrative example in Berry and Pakes (2000), which says that cost estimates will be biased toward zero if we ignore the inter-temporal effect. Investigating **a monopolist**'s pricing problem for a **single** experience good, Berry and Pakes (2000) show, through Monte Carlo simulation, that there is a negative correlation between cost variables and the inter-temporal effect that is omitted in the myopic model. The intuition for this is that since customers with experience are more likely to buy again in the future, a monopolist taking into account the benefit of future purchases would not raise prices as much as a myopic firm would in response to a cost increase. Since the cereal market is inertial overall, and an inertial good has the same demand properties as an experience good, one would also expect a negative correlation between the cost variables and the omitted inter-temporal effect. However in the context of **oligopolistic multi-product** price competition as in our case, *the effects of inertia may not be large enough to dominate the effects of marketing activities on household's brand choices*. The net effect of inertia and marketing mix is given by $\Delta_1 = \overline{\theta}_{r2r1} - \sum_{k=1,k\neq r}^{J} \overline{\theta}_{r2k1}$ (as in equation (3)). In the case of extreme inertia, in which consumers do not respond to the marketing mix, Δ_1=1 because $\overline{\theta}_{r2r1} = 1 > \sum_{k=1,k\neq r}^{J} \overline{\theta}_{r2k1} = 0$. In most cases, however, consumers do respond to marketing activities and the sign of Δ_1 suggests whether inertia effects or competitive marketing mix effects dominate. When inertia effects dominate, $\overline{\theta}_{r2r1} > \sum_{k=1,k\neq r}^{J} \overline{\theta}_{r2k1}$ and, therefore, $\Delta_1>0$; when marketing-mix effects dominate, $\overline{\theta}_{r2r1} < \sum_{k=1,k\neq r}^{J} \overline{\theta}_{r2k1}$ and, therefore, $\Delta_1<0$. In our analysis, we find that the estimate of Δ_1 is positive in some periods and negative in other periods, suggesting that neither the inertia effects nor the competitive marketing mix effects always dominates in this market. This leads to very low correlation between the inter-temporal effect and cost variables. Therefore, we do not find any significant bias in the factor cost estimates between the myopic and inter-temporal models.

Generalizability of Findings

We assess the generalizability of our conclusions using two approaches: (1) by undertaking the proposed empirical analysis on an additional product category, and (2) using simulated data, with different levels of (i) state dependence and (ii) forward looking behavior, to study whether our main conclusions still obtain.[5]

First, we replicate our empirical analysis using scanner panel data from the ketchup category. We find that the main substantive/empirical claims obtained using the cereal data generalize to the ketchup data. One difference that emerges (not surprisingly) is that the ketchup data do not reveal a variety-seeking segment.

Second, we perform an extensive simulation analysis to study whether our finding from the cereal data about the relative contribution of state dependence versus forward-looking behavior in explaining observed prices, still holds. For the simulations, we consider a 2-brand market. We vary the following attributes for the simulations: (1) Degree of asymmetry in the market's baseline preferences for firms, i.e., symmetric versus asymmetric, (2) Degree of state dependence in the market, i.e., zero order, 2 levels of inertia (moderate, high), 2 levels of variety-seeking (moderate, high), (3) Degree of forward-looking behavior of firms (under Bertrand pricing): 1-week look-ahead versus 2-week look-ahead. The above-mentioned empirical finding is supported in the simulated data, i.e., the fit improvement (over a static model without state dependence) obtained by accounting for state dependence is much greater than that obtained by accounting for inter-temporal pricing dynamics. Further, we also find that without accounting for state dependence, when it exists, we infer collusive pricing by firms, even though true pricing behavior is Bertrand pricing, which is also consistent with our finding in the cereals data.

One reason why forward-looking behavior of firms has only limited ability to explain the price variation may be because managers make decisions about wholesale prices over a long time horizon, say, for one quarter (12 weeks) at a time, also committing to these prices in advance of the quarter (in the form of a pricing contract with the retailer). Hence it is possible that the 1-week and 2-week

[5] In the paper, we simply describe the main results to conserve space. Detailed results are available in a technical appendix from the authors.

look-ahead models that we estimate are not capturing such long-horizon planning decisions.

In order to check this issue, we perform an additional simulation analysis, that is similar to the one described above, except that we simulate data such that firms engage in 1-quarter look-ahead pricing, and commit to these prices in advance. As expected, for this data, the best fitting model is a model with 1-quarter look-ahead pricing behavior. However, we continue to find that the fit improvement obtained by accounting for state dependence is much greater than the fit improvement obtained by accounting for 1-quarter look-ahead inter-temporal pricing dynamics. Further, in terms of fit, the 1-week look-ahead pricing model serves as a good approximation to the 1-quarter look-ahead pricing model. We conjecture that this is due to the fact that dynamic effects, induced by state dependence, of price promotions on market shares do not last much longer than a few weeks.

In spite of the fact that our empirical findings generalize across two packaged goods categories, we caution the reader that our results about limited forward-looking behavior on the part of firms should not be extended, without additional research, into other decision-making domains. For example, firms may take into account inter-temporal effects and, therefore, behave in a forward-looking manner in the context of inventory management systems, where computer based solutions are extensively used.

Conclusion

A rich literature in marketing, based on demand-side analyses, has documented the effects of state dependence in households' brand choices over time. In this study, we propose an empirical procedure that allows one to investigate the effects of such state dependent demand on firms' pricing decisions and the pricing equilibrium. Towards this end, we propose a structural model of demand and supply accounting for state dependence effects. Our demand model accounts for the effects of both inertia and variety seeking, allowing them to be functions of perceived inter-product similarities between brands along observed product attributes. On the supply side, we formulate and estimate an inter-temporal model of pricing that takes into account the effects of state dependence in demand. We estimate the proposed model using scanner panel data in the breakfast cereal category, and replicate the analyses in the ketchup category. Our findings are consistent between the two categories.

First, we find that omission of state dependence effects in the demand function – as has been the case in previous empirical studies on price competition -- leads to an erroneous inference about supply side behavior. Specifically, we infer price competition between firms to be tacitly collusive when, in fact, the prices arise out of Bertrand competition. Our key empirical findings related to firm pricing behavior are summarized below.

1. We find that an inter-temporal pricing model, that allows firms to look one-period ahead while making their pricing decisions, fits the data better than does a myopic pricing model (which, in turn, fits the data better than a static pricing model, as has been typically estimated in previous work). Thus firms appear to take into account the impact on future demand of setting prices for their brands in the current period. However, the gains from employing 2- or 3-period look-ahead are minimal or negative. This finding is consistent with the hypothesis of boundedly rational firms. In many scenarios involving dynamics, researchers typically assume that agents solve a dynamic program. Our results suggest that it is possible that modeling shorter-term dynamics may lead to better fit and forecasts. Given that shorter horizon dynamic programs are computationally easier to solve, this makes it easier for firms to conditionally forecast their competitors' dynamic price responses to their chosen prices.

2. A myopic model of pricing is able to explain observed prices of brands quite well, as long as we allow for firms to account for the effects of lagged demand while setting current prices for their brands. The incremental gain in explaining observed prices from additionally modeling inter-temporal effects, i.e., firms accounting for the effects of current prices on future demand for their brands, is not large. This suggests that empirical industrial organization applications can benefit significantly from using household-level data, in order to effectively estimate the effects of lagged choices on current demand, in markets where state dependence is likely to exist. Even if inter-temporal effects are ignored, the resulting myopic model seems to provide a pricing solution that is a good approximation to the true solution of the dynamic pricing game being played by firms (as discussed in the previous finding), in terms of enabling firms to obtain good conditional forecasts of their competitors' prices.

On the demand side, our study is the first to uncover variety seeking among brands (in addition to inertia) in scanner panel data (Note: Trivedi, Bass and Rao (1994) uncover variety seeking in the video rental market using survey data). One reason for previous studies' inability to estimate variety seeking could be that they aggregate SKUs to the brand-level and use the top five or so brands only in the estimation, which ends up masking households' variety-induced switching among different attribute bundles (flavors, forms etc.) of the same brand over time. In the breakfast cereal category, since SKUs are typically marketed as separate brands, we do not face this aggregation issue. By treating each of a large number of brands (SKUs) as a distinct bundle of attributes, we are able to uncover variety-seeking that is driven by (some) consumers seeking different attributes to consume from one purchase occasion to the next.

In this paper, we take an important first step in empirically estimating the responsiveness of firms' pricing decisions to state dependent demand. Data limitations, as well as computational difficulties, lead us to abstract away from many issues that should be addressed in future work.

First, to the extent that aggregate data are generally more representative (than household-level data that we use) of the marketplace, as shown by Gupta, Chintagunta, Kaul and Wittink (1996), and are also more readily available to managers for decision-making purposes, it would be interesting, but challenging, to see to what extent one could answer the questions investigated in this study using store-level or market-level demand data.

It would be interesting to study the effects of multi-market presence of the retailer. As in all previous research with local market data, we abstract away from this issue in the paper. Apart from the problem of obtaining data across multiple markets in which a retailer is present, handling the effects of multi-market retailer presence warrants solutions to a number of modeling challenges: (i) defining the geographic scope of the market for a retail chain, (ii) allowing for differences in retail competition in different retail markets etc. While investigating these effects was beyond the scope of this study, we believe that this would be a challenging and important issue for future research.

A systematic investigation of the effects of periodicity of the available data, as well as the incorporation of retailer behavior, on inference of competitive behavior can be insightful. This would require the availability of weekly data on multiple geographic markets in a single category, so that we have enough degrees of freedom to still estimate the model even after aggregating data to bi-weekly, monthly and quarterly levels. It is also possible that while retailers set their prices weekly, manufacturers may have a different periodicity of pricing. Modeling the pricing decision at the frequency at which those decisions are made in the real world would be important in future work.

In the absence of data on wholesale prices, we follow the approach developed by Sudhir (2001b) and Berto Villas-Boas (2004) in modeling manufacturer-retailer interactions. With wholesale price data, it would be possible to investigate more flexible models of interactions between manufacturers and retailers (e.g., Kadiyali et al. 2000). One possibility would be to empirically estimate a model of bargaining between manufacturers and retailers. Another interesting direction would be to investigate how state dependence will affect forward looking behavior in environments with targeted pricing (e.g., where there are personalized Catalina coupons).

Our focus in this paper is on demand dynamics that arise due to state dependence. Future research should systematically investigate how other sources of demand dynamics, such as consumer stockpiling and retailer forward-buying, affect supply-side behavior. Another source of demand side dynamics is the changes in consumers' price and promotion sensitivity over time, as a function of past promotions and advertising activities (e.g., Kopalle et al. 1999). An investigation of whether and how firms respond to such dynamic demand changes would also be of interest in future research. In past research, researchers typically either completely ignore these sources of dynamics or assume that firms solve full-fledged dynamic games. Our results suggest that it is possible that actual firm behavior may lie somewhere in between. A systematic investigation of the extent to which firms are forward looking when dealing with the several sources of demand dynamics mentioned above would be a fruitful area for future research. Lastly, one could investigate the effects of accommodating higher-order dynamics in demand (since our demand model allows for first-order lagged effects of brand choices only).

Appendix

Myopic Pricing Model

Suppose that each manufacturer, as well as the retailer, is myopic in pricing and only maximizes current-period profits. Here, we will derive the myopic pricing model under the assumptions of (1) a Bertrand-Nash game among multiple manufacturers, and (2) a Stackelberg game between the manufacturers and the single retailer.

The Retailer

The retailer takes the wholesale prices as given in the Stackelberg game, and acts like a monopolist in pricing the whole category.

At any period t, the retailer's profit maximization problem is, $\max_{p_{1t} \cdots p_{Jt}} \pi_{Rt} = \sum_{j=1}^{J} (p_{jt} - w_{jt}) s_{jt}(p)$,

where $s_j(p)$ is the market share of brand j, which is a function of the prices of all brands, M is the size of the market. The first-order condition is $\frac{\partial \pi_{Rt}}{\partial p_{j1}} = 0$, which implies, $s_{jt}(p) + \sum_{r=1}^{J} (p_{rt} - w_{rt}) \frac{\partial s_{rt}(p)}{\partial p_{j2}} = 0$. Defining $\Phi_{jrt} = -\frac{\partial s_{rt}(p)}{\partial p_{jt}}, j, r = 1, ..., J$, and using the matrix form, the retailer's margin is $p_t - w_t = \Phi_t^{-1} s_t(p)$.

The Manufacturer

Manufacturers' markups can be solved for explicitly by defining $\Phi_{jr_t} = -\frac{\partial s_{rt}(p)}{\partial p_{jt}}, j, r = 1, ..., J$, $\Omega_{jr}^{*} = \begin{cases} 1, & \text{if } \exists f : \{r, j\} \subset F_f \\ 0, & \text{otherwise} \end{cases}$ and Ω is a $J \times J$ matrix with $\Omega_{jrt} = \Omega_{jr}^{*} \times \Phi_{jrt}$. Following the same derivation as in the retailer pricing case, the manufacturer's margin is $w_t - mc_t = -(\Phi_t' G_t^{-1} \Omega_t)^{-1} s_t(p)$.

Price-Cost Margin: Summary of Myopic Model

The price-cost margin is defined as $p_t - mc = (p_t - w_t) + (w_t - mc)$. Based on the above discussion, we see that $p_t - mc = \Phi_t^{-1} s_t(p) - (\Phi_t' G_t^{-1} \Omega_t)^{-1} s_t(p)$.

One-Period Look-Ahead Pricing Model

Suppose that each manufacturer, as well as the retailer, looks ahead one-period in pricing to maximize inter-temporal profits. Here, we will derive the two-period pricing model under the assumptions of (1) a Bertrand-Nash game among multiple manufacturers, and (2) a Stackelberg game between the manufacturers and the single retailer.

The Retailer

The retailer takes the wholesale prices as given in the Stackelberg game, and acts like a monopolist in pricing the whole category.

In the one-period model, at period t $(t=1,2)$, the retailer's profit maximization problem is,

$\max_{p_{1t}\cdots p_{Jt}} \pi_{Rt} = \sum_{j=1}^{J} (p_{jt} - w_{jt}) s_{jt}(p)$, where $s_j(p)$ is the market share of brand j, which is a function of the prices of all brands, and M is the size of the market.

If the monopolist retailer, instead of being myopic, looks one-period ahead and maximizes two-period profits, the retailer's objective function will be, $V_R = \pi_{R1}(p_{j1}) + \delta \pi_{R2}(p_{j2})$, where $j = 1,...,J$.

The first-order conditions now become,

$$\frac{\partial \pi_{r1}}{\partial p_{j1}} + \delta \sum_{r=1}^{J} \frac{\partial \pi_{r2}}{\partial s_{r2}(p)} \frac{\partial s_{r2}(p)}{\partial s_{r1}(p)} \frac{\partial s_{r1}(p)}{\partial p_{j1}} = 0$$

$$\frac{\partial \pi_{r2}}{\partial p_{j2}} = 0$$

It is obvious that $\frac{\partial \pi_{R2}}{\partial s_{r2}(p)} = p_{r2} - w_{r2}$. In order to calculate $\frac{\partial s_{r2}(p)}{\partial s_{r1}(p)}$, we assume the following relationship holds between s_{r1} and s_{r2}, $s_{r2} = \theta_{r2r1} * s_{r1} + \sum_{k=1,k\neq r}^{J} \theta_{r2k1} * s_{k1}$, where, as defined in the demand models, θ_{r2r1} is the transition probability that consumers who **bought** brand r in period 1 **continue to buy** it during period 2, and θ_{r2k1} is the transition probability that consumers who **did not buy** brand r in period 1 **switch to it** during period 2.

The following relationship holds,

$$\begin{aligned} s_{r2} &= \theta_{r2r1} * s_{r1} + \sum_{k=1,k\neq r}^{J} \theta_{r2k1} * s_{k1} \\ &= \theta_{r2r1} * s_{r1} + \sum_{k=1,k\neq r}^{J} \theta_{r2k1} * (1 - s_{11} - s_{21} - \ldots - s_{r1} - \ldots - s_{k-1,1} - s_{k+1,1} - \ldots - s_{J1}) \end{aligned}$$

and therefore,

$$\frac{\partial s_{r2}(p)}{\partial s_{r1}(p)} = \theta_{r2r1} - \sum_{k=1,k\neq r}^{J} \theta_{r2k1} = \begin{cases} >0 & \text{if } SD>0 \text{ and } \theta_{r2r1} - \sum_{k=1,k\neq r}^{J} \theta_{r2k1} > 0 \\ <0 & \text{if } SD<0 \text{ and } \theta_{r2r1} - \sum_{k=1,k\neq r}^{J} \theta_{r2k1} < 0 \end{cases}$$

since $\sum_{k=1,k\neq r}^{J} \theta_{r2k1} = 1-\theta_{r2r1}$. With these two terms derived, we could now write down the retailer's price functions for both period 1 and period 2.

The detailed first-order conditions could be written as, $s_{j1}(p) + \sum_{r=1}^{J}(p_{r1}-w_{r1})\frac{\partial s_{r1}(p)}{\partial p_{j1}} + \delta\sum_{r=1}^{J}(p_{r2}-w_{r2})(\Delta_1)\frac{\partial s_{r1}(p)}{\partial p_{j1}} = 0$ for period 1; and

$s_{j2}(p) + \sum_{r=1}^{J}(p_{r2}-w_{r2})\frac{\partial s_{r2}(p)}{\partial p_{j2}} = 0$ for period 2, where $\Delta_1 = \theta_{r2r1} - \sum_{k=1,k\neq r}^{J}\theta_{r2k1}$.

Defining $\Phi_{jrt} = -\frac{\partial s_r(p)}{\partial p_j}$, $j,r=1,...,J$, and using the matrix form, the retailer's **second-period pricing function** could be written as $p_2 - w_2 = \Phi_2^{-1}s_2(p)$ and **the retailer's first period pricing function** could be written as $p_1 - w_1 = \Phi_1^{-1}\left\{s_1(p) - \Phi_1\left[\delta\Phi_2^{-1}s_2(p)(\Delta_1)\right]\right\}$.

The Manufacturer

Manufacturers' markups can be solved for explicitly by defining $\Phi_{jr_t} = -\frac{\partial s_{rt}(p)}{\partial p_{jt}}$, $j,r=1,...,J$,

$\Omega^*_{jr} = \begin{cases} 1, & \text{if } \exists f : \{r,j\} \subset F_f \\ 0, & \text{otherwise} \end{cases}$ and Ω is a $J \times J$ matrix with $\Omega_{jrt} = \Omega^*_{jr} \times \Phi_{jrt}$. Following the same derivation as in the retailer pricing case, **in period 2, the** manufacturer's margin is $w_2 - mc = -(\Phi_2' G_2^{-1}\Omega_2)^{-1}s_2(p)$ and **in the first-period, the manufacturer's** margin is, $w_1 - mc = \Omega_1^{-1}\left\{s_1(p) + \Omega_1\left[\delta(\Phi_2' G_2^{-1}\Omega_2)^{-1}s_2(p)(\Delta_1)\right]\right\}$

Price-Cost Margin: Summary of the One-Period Look Ahead Model

The price-cost margin is defined as $p_t - mc = (p_t - w_t) + (w_t - mc)$. Based on the above discussion, we see that

$$
p_t - mc = \begin{cases} \Phi_1^{-1} s_1(p) - \delta \Phi_2^{-1} s_2(p)\left(\Delta_1\right) + \Omega_1^{-1} s_1(p) + \delta(\Phi_2' G_2^{-1} \Omega_2)^{-1} s_2(p)\left(\Delta_1\right) \\ = PCM^{Myopic}(t=1) - \delta \Delta_1 PCM^{Myopic}(t=2) & \text{when } t = 1 \\ \\ \Phi_2^{-1} s_2(p) - (\Phi_2' G_2^{-1} \Omega_2)^{-1} s_2(p) = PCM^{Myopic}(t=2) & \text{when } t = 2 \end{cases} \quad (9)
$$

where $\Delta_1 = \theta_{r2 \cdot r1} - \sum_{k=1, k \neq r}^{J} \theta_{r2 \cdot k1}$

References

Ailawadi, Kusum L., Praveen K. Kopalle and Scott A. Neslin (2005), "Predicting Competitive Response to a Major Policy Change: Combining Game-Theoretic and Empirical Analyses", *Marketing Science,* 24, 1, 1-12.

Andrews, Rick and T.C. Srinivasan (1995), “Studying Consideration Effects in Empirical Choice Models Using Scanner Panel Data,” *Journal of Marketing Research*, 32, 1, 30-41.

Berry, Steve and Ariel Pakes (2000), “Estimation from the First Order Conditions for Dynamic Controls”, Working Paper, Department of Economics, Yale University.

Besanko, David, Sachin Gupta and Dipak Jain (1998), "Logit Demand Estimation under Competitive Pricing Behavior: An Equilibrium Framework", *Management Science,* 44, 11, 1533-1547.

Camerer, Colin F., Teck Hua Ho, and Juin-kuan Chong (2004), "A Cognitive Hierarchy Model of Games", *The Quarterly Journal of Economics,* 119, 3, 861-898.

Chakravarti, Dipankar, Andrew Mitchell, and Richard Staelin (1979), "Judgment Based Marketing Decision Models: An Experimental Investigation of the Decision Calculus Approach", *Management Science,* 25, 3, 251-262.

Erdem, Tulin (1996), "A Dynamic Analysis of Market Structure Based on Panel Data", *Marketing Science,* 15, 4, 359-378.

Erdem, Tulin, Susumo Imai, and Michael P. Keane (2003), “A Model of Consumer Brand and Quantity Choice Dynamics Under Price Uncertainty”, *Quantitative Marketing and Economics*, 1, 1, 5-64.

Fader, Peter S., Bruce G.S. Hardie, (1996), "Modeling Consumer Choice Among SKUs", *Journal of Marketing Research,* 33, 4, 442-452.

Freimer, Marshall, and Dan Horsky (2004), “Try It, You Will Like It — Does Consumer Learning Lead to Competitive Price Promotion?” Working Paper, Simon School of Business, University of Rochester.

Genesove, David and Wallace P. Mullin (1998) “Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914” *The Rand Journal of Economics*, 29, 2, 355-377.

Guadagni, Peter E. and John D.C. Little (1983), "A Logit Model of Brand Choice Calibrated On Scanner Data," *Marketing Science*, 2, 2, 203-238.

Hausman, Jerry A. (1996) "Valuation of New Goods Under Perfect and Imperfect Competition," in *The Economics of New Products*, T. F. Bresnahan and R. J. Gordon, eds. Chicago: University of

Chicago Press, 1996.

Hausman, Jerry A. and McFadden, Daniel (1984), "Specification Tests for the Multinomial Logit Model", *Econometrica,* 52, 5, 1219-1240.

Ho, Teck H. and Juin-Kuan Chong (2003), "A Parsimonious Model of Stockkeeping-Unit Choice", *Journal of Marketing Research,* 40, 3, 351-365.

Ho, Teck H., Colin Camerer and Noah Lim (2006), "Modeling the Psychology of Consumer and Firm Behavior Using Behavioral Economics", *Journal of Marketing Research,* forthcoming.

Kadiyali, Vrinda, K. Sudhir and Vithala R. Rao (2001), "Structural Analysis of Competitive Behavior: New Empirical Industrial Organization Methods in Marketing", *International Journal of Research in Marketing,* 18, 2, 161-186.

Kadiyali, Vrinda (1996), "Entry, its Deterrence, and its Accommodation: A Study of the U.S. Photographic Film Industry", *The Rand Journal of Economics,* 27, 3, 452-478.

Kahn, Barbara E., Manohar Kalwani and Donald Morrison (1986) "Measuring Variety-Seeking and Reinforcement Behaviors Using Panel Data," *Journal of Marketing Research*, 23, 1, 89-100.

Kamakura, Wagner A., and Gary J. Russell (1989), "A Probabilistic Choice Model For Market Segmentation And Elasticity Structure", *JMR, Journal of Marketing Research,* 26, 4, 379-390.

Keane, Michael P. (1997), "Modeling Heterogeneity and State Dependence in Consumer Choice Behavior", *Journal of Business and Economic Statistics,* 15, 3, 310-327.

Klemperer, Paul (1987), "Markets with Consumer Switching Costs", *The Quarterly Journal of Economics,* 102, 2, 375-394.

Kopalle, Praveen K., Carl F. Mela, and Lawrence Marsh (1999), "The Dynamic Effect of Discounting on Sales: Empirical Analysis and Normative Pricing Implications", *Marketing Science,* 18, 3, 317-332.

Lattin, James M. and Leigh McAlister (1985), "Using a Variety-Seeking Model to Identify Substitute and Complementary Relationships Among Competing Products", *Journal of Marketing Research,* 22, 3, 330-339.

Lee, Jack, K. Sudhir and Joel Steckel (2002), "A Multiple Ideal Point Model: Capturing Multiple Preference Effects from Within an Ideal Point Framework," *Journal of Marketing Research,* 39, 1, 73-86.

McAlister, Leigh and Edgar Pessemier (1982), "Variety Seeking Behavior: An Interdisciplinary Review," *Journal of Consumer Research,* 9, 3, 311-322.

McFadden, Daniel (1980), “Econometric Models of Probabilistic Choice”, in Structural Analysis of Discrete Data with Econometric Applications, ed. By C. Manski and D. McFadden, Cambridge, Mass: MIT Press.

Mcintyre, Shelby H. (1982), "An Experimental Study of the Impact of Judgment-Based Marketing Models", *Management Science,* 28, 1, 17-33.

Meyer, Robert J. and J. Wesley Hutchinson, (2001), Chapter 3: Bumbling Geniuses: The Power of Everyday Reasoning in Multistage Decision Making. Wharton on Making Decisions, Wiley.

Nevo, Aviv (2001), "Measuring Market Power in the Ready-to-Eat Cereal Industry", *Econometrica,* 69, 2, 307-342.

Newey, West K. and Daniel McFadden, (1994), Handbook of Econometrics Vol. IV, Chapter 36, Large Sample Estimation and Hypothesis Testing, New York, Elsevier.

Roy, Abhik, Dominique M. Hanssens, and Jagmohan S. Raju (1994), "Competitive Pricing by a Price Leader", *Management Science,* 40, 7, 809-823.

Seetharaman, P.B. (2004), "Modeling Multiple Sources of State Dependence in Random Utility Models: A Distributed Lag Approach", *Marketing Science,* 23, 2, 263-271.

Seetharaman, P.B. and Pradeep K. Chintagunta, (1998), "A Model of Inertia and Variety-Seeking with Marketing Variables", *International Journal of Research in Marketing,* 15, 1, 1-17.

Seetharaman, P.B., Andrew Ainslie, and Pradeep K. Chintagunta, (1999), "Investigating Household State Dependence Effects Across Categories", *Journal of Marketing Research,* 36, 4, 488-500.

Sudhir, K. (2001a), "Competitive Pricing Behavior in the Auto Market: A Structural Analysis", *Marketing Science,* 20, 1, 42-60.

Sudhir, K. (2001b), "Structural Analysis of Manufacturer Pricing in the Presence of a Strategic Retailer", *Marketing Science,* 20, 3, 244-264.

Trivedi, Minakshi, Frank M. Bass, and Ram C. Rao, (1994), "A Model of Stochastic Variety-Seeking", *Marketing Science,* 13, 3, 274-297.

Villas-Boas, J. Miguel (2004), "Consumer Learning, Brand Loyalty, and Competition", *Marketing Science,* 23, 1, 134-145.

Villas-Boas, J. Miguel and Russell S. Winer, (1999), "Endogeneity in Brand Choice Models", *Management Science,* 45, 10, 1324-1338.

Villas-Boas, J. Miguel and Ying Zhao, (2005), "Retailer, Manufacturers, and Individual Consumers: Modeling the Supply Side in the Ketchup Marketplace", *Journal of Marketing Research,* 42, 1, 83-95.

Villas-Boas, Sofia B. (2004), “Vertical Contracts between Manufacturers and Retailers: An Empirical Analysis”, Working Paper, University of California at Berkeley.

Walters, Rockney (1991),"Assessing the Impact of Retail Price Promotions on Product Substitution, Complementary Purchase, and Interstore Sales Displacement", *Journal of Marketing,* 55, 2, 17-28.

Figure 1: A Schematic Model of the Market

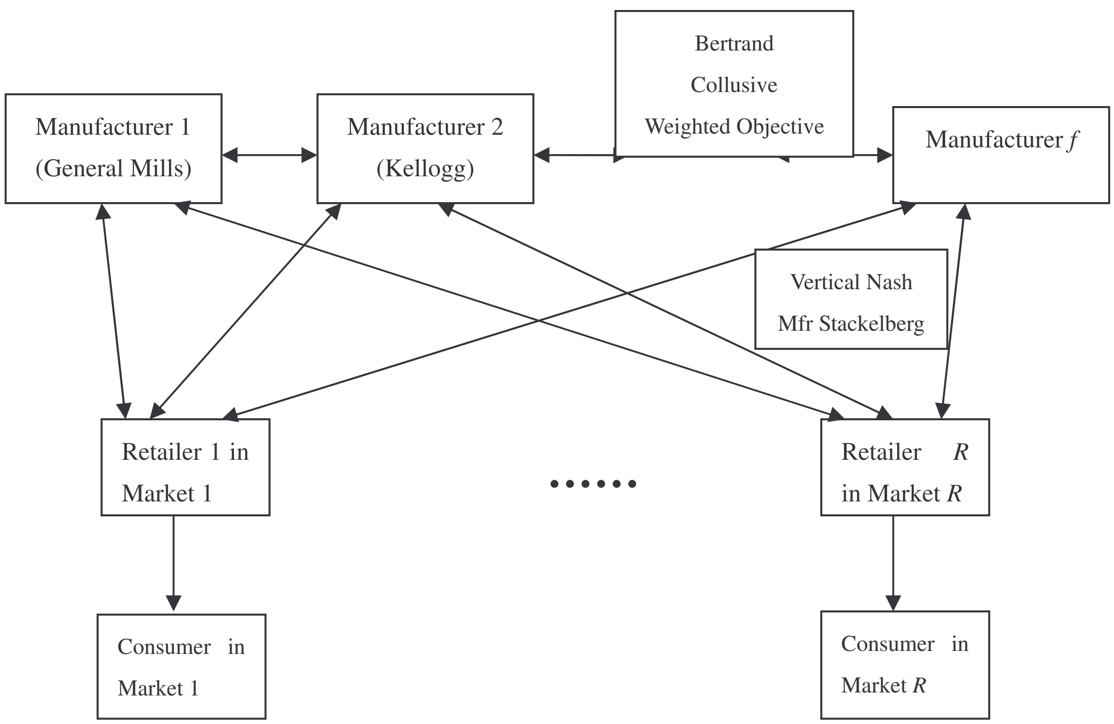

Table 1: Demand-Side Estimation: Fit Results

Model Specification	Multinomial Logit no State-Dependence	State-Dependence with Lag Dummies	State-Dependence with Similarity Terms
1. Without Heterogeneity			
Number of Parameters	8	13	18
Price Coefficients	-25.50	-23.23	-23.09
State-dependence Estimates		2.87	3.15
Log-likelihood	-146949	-137997	-137407
AIC	293914	276020	274850
BIC	293937	276057	274901
2. With Heterogeneity			
(2-Support Case)			
Number of Parameters	17	27	32
Price Coefficients	-91.28	-74.45	-66.35
	-3.25	-.49	.32
State-dependence Estimates		1.60	.01
		3.70	4.55
Log-likelihood	-141525	-132171	-131388
AIC	283084	264396	262840
BIC	283034	264340	262778
2. With Heterogeneity			
(3-Support Case)			
Number of Parameters	26	41	46
Price Coefficients	-90.65	-89.87	-85.46
	-2.19	-1.92	-1.37
	-2.98	-2.05	-3.50
State-dependence Estimates		.6784	.6803
		4.33	4.83
		-1.42	-1.46
Log-likelihood	-139959	-130756	-129930
AIC	279970	261594	259952
BIC	279918	261515	259863
Number of Observations:	68,312		

Table 2: Demand-Side Estimation:

Parameter Estimates for 3-Support Model With State-Dependence

	No Endogeneity Correction			With Endogeneity Correction		
Variables	Support 1 Estimates	Support 2 Estimates	Support 3 Estimates	Support 1 Estimates	Support 2 Estimates	Support 3 Estimates
Constant	**-6.85**	**.31**	**.02**	**-7.00**	**.25**	**.51**
	(.26)	(1.23)	(.17)	(.30)	(1.24)	(.17)
Attributes						
Sweetened	-.12	**-.33**	**.22**	-.03	**-.41**	**.23**
	(.07)	(.04)	(.06)	(.08)	(.04)	(.06)
Fiber	**.28**	**.08**	**-3.08**	**.42**	**.09**	**-2.84**
	(.07)	(.04)	(.06)	(.08)	(.04)	(.06)
Fruit-nuts	**.24**	.02	**.60**	**.14**	**-.08**	**.40**
	(.06)	(.04)	(.04)	(.06)	(.04)	(.04)
Marketing Mix						
Price	**-3.50**	**-1.37**	**-85.46**	**-4.08**	**-2.06**	**-87.24**
	(.90)	(.42)	(1.16)	(1.02)	(.48)	(1.18)
Feature	**.27**	**.92**	**-6.62**	**.36**	**.90**	**-6.09**
	(.16)	(.10)	(.16)	(.16)	(.10)	(.16)
Display	**.28**	**.22**	**.24**	**.25**	**.35**	**.21**
	(.17)	(.09)	(.08)	(.17)	(.10)	(.08)
Inventory	**.96**	-.04	.01	**.99**	-.04	.01
	(.06)	(.10)	(.00)	(.07)	(.11)	(.00)
Demographics						
Intercept	**-1.35**	**.77**	**.76**	**-1.23**	**.02**	**.87**
	(.39)	(.05)	(.14)	(.41)	(.05)	(.12)
Family size	**.25**	**-.16**	**.27**	**.23**	**-.16**	**.24**
	(.12)	(.03)	(.06)	(.13)	(.03)	(.05)
Income	**.76**	**-.10**	**.33**	**.82**	**-.12**	**.32**
	(.06)	(.01)	(.03)	(.07)	(.01)	(.02)
Age	**-.72**	**.05**	**-.38**	**-.59**	**.07**	**-.38**
	(.13)	(.03)	(.06)	(.14)	(.03)	(.05)
Child age	**-.67**	**-.27**	**-.33**	**-.70**	**-.31**	**-.32**
	(.09)	(.03)	(.04)	(.10)	(.03)	(.03)
State-dependence Estimates						
Mean	-1.46	4.83	.68	-1.35	5.08	.79
Standard error	.33	.09	.11	.34	.10	.10
Similarity						
Ingredient	.09			.09		
	(1.58)			(1.57)		
Sweetened	**.17**			**.16**		
	(.01)			(.01)		
Fiber	.09			.09		
	(1.58)			(1.57)		
Fruit-nuts	**.08**			**.09**		
	(.01)			(.01)		
Product type	**.11**			**.11**		
	(.01)			(.01)		
Segments						
Size	29.17%	17.29%	53.54%	27.33%	16.29%	56.38%
	(.01)	(.00)		(.01)	(.00)	
Number of Parameters	46			66		
Number of Observations	68,312			68,312		
LL	-129,930			-56,529		

Table 3: Supply-Side Estimation: Fit Results

Manufacturer Interaction	Retailer Objective Manufacturer-Retailer Interaction	Log-Likelihoods (Vuong Test Statistics)		
		Static Pricing	Myopic Pricing	One-Period Look Ahead Pricing
		Without State-Dependence	With State-Dependence	With State-Dependence
Tacit Collusion	**Category Profit Maximization Manufacturer Vertical Nash**	31033.2 (1.87*)	31484.2 (4.76**)	33811.1 (3.84**)
	Category Profit Maximization Manufacturer Stackelberg	**31308.4** (—)	32793.8 (5.31**)	31600.4 (3.99**)
Bertrand Competition	**Category Profit Maximization Manufacturer Vertical Nash**	30865.4 (3.97**)	33384.2 (2.39*)	34115.3 (1.95*)
	Category Profit Maximization Manufacturer Stackelberg	30390.6 (5.13**)	**33923.4** (—)	**34455.9** (—)

(—): the best fitting model

***: p <.1; ** p <.05**

Significance level in accepting the best fitting model and rejecting the alternative model based on Vuong test statistics

Table 4: Price Fitting Errors

	Mean Squared Errors (MSE) of Fitted Prices	Variance of MSE	Reduction in (%)
Static Model without State-dependence	.053	.23	
Myopic Model with State-dependence	.041	.193	Compared with Static No SD 22.34%
One-period look ahead Model with State-dependence	.041	.19	Compared with Static No SD 23.67% Compared with Myopic With SD 1.71%
Two-period look ahead Model with State-dependence	.046	.20	Compared with Static No SD 13.59% Compared with One-period With SD -11.70%

Table 5: Supply-Side Estimation
One-period Look-Ahead Pricing Models Using Demand Estimates With State-Dependence

		Bertrand (Vertical Nash)			Collusive (Vertical Nash)		
Parameters		**Estimates**	**S.E.**	**t-stat**	**Estimates**	**S.E.**	**t-stat**
Manufacturer Dummies	Nabisco	**.0156**	.00	18.90	**.0152**	.00	16.20
	General Mills	**-.0225**	.00	-10.73	**-.0248**	.00	-11.35
	Post	**.0035**	.00	2.93	**.0028**	.00	2.12
	Quaker	**.0111**	.00	12.00	**.0102**	.00	9.63
	Kellogg	**.0103**	.00	13.56	**.0124**	.00	14.61
Store Dummies	Store 2 Dummy	**-.0097**	.00	-21.09	**-.0110**	.00	-21.56
	Store 3 Dummy	**-.0057**	.00	-11.40	**-.0059**	.00	-11.23
Cost	Constant	**.1726**	.00	193.04	**.1695**	.00	171.56
	Ingredient	.0003	.00	.53	.0000	.00	-.06
	Sugar	**-.0033**	.00	-3.58	**-.0046**	.00	-4.58
	Fruit	**-.0440**	.00	-81.64	**-.0442**	.00	-73.49
	Fiber	**.0001**	.00	8.15	**.0001**	.00	7.84
	Wage	.0000	.00	.55	.0000	.00	-.73
	Gasoline	.0000	.00	-.69	-.0001	.00	-.92
	Packaging	**.0009**	.00	4.04	**.0004**	.00	1.91
LL				34115.3			33811.1
		Bertrand (Stackelberg)			**Collusive (Stackelberg)**		
Parameters		**Estimates**	**S.E.**	**t-stat**	**Estimates**	**S.E.**	**t-stat**
Manufacturer Dummies	Nabisco	**.0162**	.00	19.14	**.0159**	.00	8.27
	General Mills	**-.0253**	.00	-11.84	**-.0362**	.00	-13.09
	Post	**.0037**	.00	3.06	**.0037**	.00	1.77
	Quaker	**.0112**	.00	11.94	**.0080**	.00	4.86
	Kellogg	**.0098**	.00	12.56	**.0197**	.00	12.22
Store Dummies	Store 2 Dummy	**-.0097**	.00	-21.04	**-.0127**	.00	-23.83
	Store 3 Dummy	**-.0058**	.00	-11.63	**-.0043**	.00	-7.20
Cost	Constant	**.1727**	.00	189.74	**.1598**	.00	95.51
	Ingredient	.0003	.00	.51	**-.0013**	.00	-2.19
	Sugar	**-.0034**	.00	-3.71	**-.0086**	.00	-9.06
	Fruit	**-.0447**	.00	-81.46	**-.0459**	.00	-77.71
	Fiber	**.0001**	.00	8.14	**.0001**	.00	7.63
	Wage	.0000	.00	-.60	.0000	.00	-.40
	Gasoline	.0000	.00	-.83	-.0001	.00	-.97
	Paper	**.0010**	.00	4.46	**.0011**	.00	5.38
LL				34455.9			31600.4
Number of Parameters	113						
Number of Observations	27,040						

Note 1: For identification, one of the manufacturer dummies (Ralston) and one of the store dummies are set to 0.

Note 2: Store 2 and 3 on average carry a smaller number of brands every week.